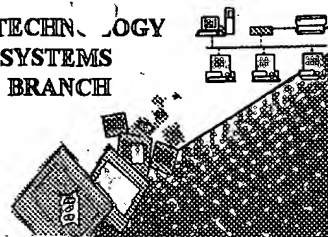


RAW SEQUENCE LISTING
ERROR REPORT

BIOTECHNOLOGY
SYSTEMS
BRANCH



File Log

1600

1636

#14

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/689,343c
Source: 1600
Date Processed by STIC: 7/19/02

RECEIVED

AUG 14 2002

TECH CENTER 1600/2900

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/efc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

RECEIVED

AUG 14 2002

TECH CENTER 1600/2900



1600

Does Not Comply
Corrected Diskette Needed

RAW SEQUENCE LISTING

DATE: 07/19/2002

Error on p. 5

PATENT APPLICATION: US/09/689,343C

TIME: 14:59:20

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\07192002\I689343C.raw

3 <110> APPLICANT: Vaisvila, Romualdus
 4 Morgan, Richard D.
 5 Kucera, Rebecca B.
 6 Claus, Toby B.
 7 Raleigh, Elisabeth A.
 9 <120> TITLE OF INVENTION: Method For Cloning And Producing The MseI Restriction
 Endonuclease
 11 <130> FILE REFERENCE: NEB-181
 13 <140> CURRENT APPLICATION NUMBER: US 09/689,343C
 14 <141> CURRENT FILING DATE: 2000-10-12
 16 <160> NUMBER OF SEQ ID NOS: 9
 18 <170> SOFTWARE: PatentIn version 3.1
 20 <210> SEQ ID NO: 1
 21 <211> LENGTH: 903
 22 <212> TYPE: DNA
 23 <213> ORGANISM: Micrococcus sp.
 25 <220> FEATURE:
 26 <221> NAME/KEY: CDS
 27 <222> LOCATION: (1)..(900)
 28 <223> OTHER INFORMATION:
 30 <400> SEQUENCE: 1
 32 atg cct atc tcg acc gtc tgg acg ccg gac gga gac gac ctc atc gtg 48
 33 Met Pro Ile Ser Thr Val Trp Thr Pro Asp Gly Asp Asp Leu Ile Val
 34 1 5 10 15
 36 gag gcg gac aac ctc gat ttc att caa acg ctc ccc gac gcg agc ttc 96
 37 Glu Ala Asp Asn Leu Asp Phe Ile Gln Thr Leu Pro Asp Ala Ser Phe
 38 20 25 30
 40 cga atg atc tac atc gat ccg ccg ttc aac aca ggg cga acg cag cgg 144
 41 Arg Met Ile Tyr Ile Asp Pro Pro Phe Asn Thr Gly Arg Thr Gln Arg
 42 35 40 45
 44 ctt cag tcg ctc aag acg acc cgc tcg gtc aca ggg tcg cga gtc ggc 192
 45 Leu Gln Ser Leu Lys Thr Thr Arg Ser Val Thr Gly Ser Arg Val Gly
 46 50 55 60
 48 ttc aaa ggc cag acg tac gac acg gtc aag agc act ctg cac tcg tat 240
 49 Phe Lys Gly Gln Thr Tyr Asp Thr Val Lys Ser Thr Leu His Ser Tyr
 50 65 70 75 80
 52 gac gac gct ttc acc gac tat tgg tcg ttc ctc gaa ccg cgt ctc ctg 288
 53 Asp Asp Ala Phe Thr Asp Tyr Trp Ser Phe Leu Glu Pro Arg Leu Leu
 54 85 90 95
 56 gag gct tgg cgg ttg ctc acc cct gac ggc gcg ctc tat ctt cat ctg 336
 57 Glu Ala Trp Arg Leu Leu Thr Pro Asp Gly Ala Leu Tyr Leu His Leu
 58 100 105 110
 60 gat tac cgc gag gtt cac tac gcc aag gtc gtc ctc gac gcg atg ttc 384
 61 Asp Tyr Arg Glu Val His Tyr Ala Lys Val Val Leu Asp Ala Met Phe

RAW SEQUENCE LISTING

DATE: 07/19/2002

PATENT APPLICATION: US/09/689,343C

TIME: 14:59:20

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\07192002\I689343C.raw

```

62          115          120          125
64 gga cgc gaa agc ttc ctg aac gag ctg atc tgg gcg tac gac tac ggc      432
65 Gly Arg Glu Ser Phe Leu Asn Glu Leu Ile Trp Ala Tyr Asp Tyr Gly
66          130          135          140
68 gcg cgc tcg aag agc aag tgg ccc acc aag cac gac aac atc ctc gtg      480
69 Ala Arg Ser Lys Ser Lys Trp Pro Thr Lys His Asp Asn Ile Leu Val
70 145          150          155          160
72 tat gtg aag gac ccg aac aac tac gtc tgg aac ggt cag gat gta gat      528
73 Tyr Val Lys Asp Pro Asn Asn Tyr Val Trp Asn Gly Gln Asp Val Asp
74          165          170          175
76 cgc gag ccc tac atg gcg ccc ggg ctc gtt aca ccc gag aag gta gcg      576
77 Arg Glu Pro Tyr Met Ala Pro Gly Leu Val Thr Pro Glu Lys Val Ala
78          180          185          190
80 ctt ggc aag ctg ccc acc gac gtc tgg tgg cac aca atc gtt ccg cct      624
81 Leu Gly Lys Leu Pro Thr Asp Val Trp Trp His Thr Ile Val Pro Pro
82          195          200          205
84 gcg agc aaa gag cgc acc ggg tac gcg aca cag aag ccg gtc ggc atc      672
85 Ala Ser Lys Glu Arg Thr Gly Tyr Ala Thr Gln Lys Pro Val Gly Ile
86          210          215          220
88 atc cgt cgc atg att cag gcg agc agc aat gaa ggc gac tgg gtt ctg      720
89 Ile Arg Arg Met Ile Gln Ala Ser Ser Asn Glu Gly Asp Trp Val Leu
90 225          230          235          240
92 gat ttc ttc gct ggt agt ggg acg acc ggc gcc gcg gcc cgc cag ctc      768
93 Asp Phe Phe Ala Gly Ser Gly Thr Thr Gly Ala Ala Ala Arg Gln Leu
94          245          250          255
96 gga cgc cgt ttt gtg ctc gta gac gtc aac cca gaa gca atc gcg gta      816
97 Gly Arg Arg Phe Val Leu Val Asp Val Asn Pro Glu Ala Ile Ala Val
98          260          265          270
100 atg gca aaa cgg ttg gat gac ggg gca ttg gac acc agc gtg acg atc      864
101 Met Ala Lys Arg Leu Asp Asp Gly Ala Leu Asp Thr Ser Val Thr Ile
102          275          280          285
104 gtg cag act ccc cag agt gac cca cga acc gac gga tga      903
105 Val Gln Thr Pro Gln Ser Asp Pro Arg Thr Asp Gly
106          290          295          300
109 <210> SEQ ID NO: 2
110 <211> LENGTH: 300
111 <212> TYPE: PRT
112 <213> ORGANISM: Micrococcus sp.
114 <400> SEQUENCE: 2
116 Met Pro Ile Ser Thr Val Trp Thr Pro Asp Gly Asp Asp Leu Ile Val
117 1          5          10          15
119 Glu Ala Asp Asn Leu Asp Phe Ile Gln Thr Leu Pro Asp Ala Ser Phe
120          20          25          30
122 Arg Met Ile Tyr Ile Asp Pro Pro Phe Asn Thr Gly Arg Thr Gln Arg
123          35          40          45
125 Leu Gln Ser Leu Lys Thr Thr Arg Ser Val Thr Gly Ser Arg Val Gly
126          50          55          60
128 Phe Lys Gly Gln Thr Tyr Asp Thr Val Lys Ser Thr Leu His Ser Tyr
129 65          70          75          80

```

RAW SEQUENCE LISTING

DATE: 07/19/2002

PATENT APPLICATION: US/09/689,343C

TIME: 14:59:20

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\07192002\I689343C.raw

```

131 Asp Asp Ala Phe Thr Asp Tyr Trp Ser Phe Leu Glu Pro Arg Leu Leu
132      85      90      95
134 Glu Ala Trp Arg Leu Leu Thr Pro Asp Gly Ala Leu Tyr Leu His Leu
135      100      105      110
137 Asp Tyr Arg Glu Val His Tyr Ala Lys Val Val Leu Asp Ala Met Phe
138      115      120      125
140 Gly Arg Glu Ser Phe Leu Asn Glu Leu Ile Trp Ala Tyr Asp Tyr Gly
141      130      135      140
143 Ala Arg Ser Lys Ser Lys Trp Pro Thr Lys His Asp Asn Ile Leu Val
144 145      150      155      160
146 Tyr Val Lys Asp Pro Asn Asn Tyr Val Trp Asn Gly Gln Asp Val Asp
147      165      170      175
149 Arg Glu Pro Tyr Met Ala Pro Gly Leu Val Thr Pro Glu Lys Val Ala
150      180      185      190
152 Leu Gly Lys Leu Pro Thr Asp Val Trp Trp His Thr Ile Val Pro Pro
153      195      200      205
155 Ala Ser Lys Glu Arg Thr Gly Tyr Ala Thr Gln Lys Pro Val Gly Ile
156      210      215      220
158 Ile Arg Arg Met Ile Gln Ala Ser Ser Asn Glu Gly Asp Trp Val Leu
159 225      230      235      240
161 Asp Phe Phe Ala Gly Ser Gly Thr Thr Gly Ala Ala Ala Arg Gln Leu
162      245      250      255
164 Gly Arg Arg Phe Val Leu Val Asp Val Asn Pro Glu Ala Ile Ala Val
165      260      265      270
167 Met Ala Lys Arg Leu Asp Asp Gly Ala Leu Asp Thr Ser Val Thr Ile
168      275      280      285
170 Val Gln Thr Pro Gln Ser Asp Pro Arg Thr Asp Gly
171      290      295      300
173 <210> SEQ ID NO: 3
174 <211> LENGTH: 1236
175 <212> TYPE: DNA
176 <213> ORGANISM: Unknown
178 <220> FEATURE:
179 <223> OTHER INFORMATION: Environmental DNA
181 <220> FEATURE:
182 <221> NAME/KEY: CDS
183 <222> LOCATION: (1)..(1233)
184 <223> OTHER INFORMATION:
186 <220> FEATURE:
187 <221> NAME/KEY: misc_feature
188 <222> LOCATION: (198)..(198)
189 <223> OTHER INFORMATION: Xaa = any amino acid
191 <220> FEATURE:
192 <221> NAME/KEY: misc_feature
193 <222> LOCATION: (594)..(594)
194 <223> OTHER INFORMATION: N= G, A, C or T
196 <400> SEQUENCE: 3
198 atg cct aca ctg gat tgg ccc ggt aaa cag tta agc ttc cca cca gct
199 Met Pro Thr Leu Asp Trp Pro Gly Lys Gln Leu Ser Phe Pro Pro Ala

```

48

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/689,343C

DATE: 07/19/2002

TIME: 14:59:20

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\07192002\I689343C.raw

200	1		5		10		15		
202	acc	tcc	ttg	cat	ctg	gag	agt	gtg	gtc
203	Thr	Ser	Leu	His	Leu	Glu	Ser	Val	Val
204				20				25	
206	cct	aat	cgt	ctg	att	tgg	gcg	gac	aac
207	Pro	Asn	Arg	Leu	Ile	Trp	Ala	Asp	Asn
208			35					40	
210	ttg	gcc	gaa	tat	gaa	ggg	aaa	atc	gat
211	Leu	Ala	Glu	Tyr	Glu	Gly	Lys	Ile	Asp
212		50					55		60
214	ttt	ttt	acg	gat	cgt	act	tat	gcg	gcg
215	Phe	Phe	Thr	Asp	Arg	Thr	Tyr	Ala	Ala
216	65				70			75	
218	tcg	cgt	cgt	cca	caa	acc	tgg	cag	ctt
219	Ser	Arg	Arg	Pro	Gln	Thr	Trp	Gln	Leu
220				85				90	
222	tgg	aag	gat	tta	gat	gaa	tac	ctg	gac
223	Trp	Lys	Asp	Leu	Asp	Glu	Tyr	Leu	Asp
224			100					105	
226	ctg	atg	tat	cga	ctg	ctg	gca	cca	cac
227	Leu	Met	Tyr	Arg	Leu	Leu	Ala	Pro	His
228			115					120	
230	gac	tgg	cac	gcc	aat	gcc	tac	gta	cgt
231	Asp	Trp	His	Ala	Asn	Ala	Tyr	Val	Arg
232		130						135	
234	ggg	cga	cag	cgg	ttt	ctc	aac	gag	atc
235	Gly	Arg	Gln	Arg	Phe	Leu	Asn	Glu	Ile
236	145				150			155	
238	tca	gcc	atc	cga	cgc	gcc	ttc	aag	cgc
239	Ser	Ala	Ile	Arg	Arg	Ala	Phe	Lys	Arg
240				165				170	
242	tat	gtg	aaa	ggg	gaa	aac	tat	aca	ttc
243	Tyr	Val	Lys	Gly	Glu	Asn	Tyr	Thr	Phe
244			180					185	
W--> 246	cct	tac	cat	ccg	agc	acn	cat	aag	acc
W--> 247	Pro	Tyr	His	Pro	Ser	Xaa	His	Lys	Thr
248			195					200	
250	ggc	ttt	ggg	aag	gtg	ccg	gat	ctg	cag
251	Gly	Phe	Gly	Lys	Val	Pro	Asp	Leu	Gln
252		210						215	
254	tgg	tgg	tat	ttt	ccg	gtc	gtg	gcc	cgt
255	Trp	Trp	Tyr	Phe	Pro	Val	Val	Ala	Arg
256	225				230			235	
258	tat	ccg	act	caa	aag	cct	caa	gcc	ttg
259	Tyr	Pro	Thr	Gln	Lys	Pro	Gln	Ala	Leu
260			245					250	
262	tcc	tcg	aac	gca	ggc	gat	ctg	gtg	gca
263	Ser	Ser	Asn	Ala	Gly	Asp	Leu	Val	Ala
264			260					265	

RAW SEQUENCE LISTING

DATE: 07/19/2002

PATENT APPLICATION: US/09/689,343C

TIME: 14:59:21

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\07192002\I689343C.raw

```

266 aca acc gct gtg gtg gca gcc cgt ctg gga cgg cgc ttc ctg gtc aac      864
267 Thr Thr Ala Val Val Ala Ala Arg Leu Gly Arg Arg Phe Leu Val Asn
268      275      280      285
270 gat gca agc tgg cgc gcc gtt cat gtg aca cgc aca cgc ttg cta cgc      912
271 Asp Ala Ser Trp Arg Ala Val His Val Thr Arg Thr Arg Leu Leu Arg
272      290      295      300
274 gag gga gta agt ttc act ttt gaa cgc cag gaa act ttt act cta cct      960
275 Glu Gly Val Ser Phe Thr Phe Glu Arg Gln Glu Thr Phe Thr Leu Pro
276 305      310      315      320
278 atc cag cca ctt cca cca gat tgg ttg atc atc gcc gag gag cag att      1008
279 Ile Gln Pro Leu Pro Pro Asp Trp Leu Ile Ile Ala Glu Glu Gln Ile
280      325      330      335
282 cgc ctc caa gca ccc ttt ctc gta gat ttt tgg gaa gtg gac gat caa      1056
283 Arg Leu Gln Ala Pro Phe Leu Val Asp Phe Trp Glu Val Asp Asp Gln
284      340      345      350
286 tgg gat ggc aaa atc ttc cgc agc cgt cat caa ggc tta cgc tcc cgc      1104
287 Trp Asp Gly Lys Ile Phe Arg Ser Arg His Gln Gly Leu Arg Ser Arg
288      355      360      365
290 ctt cag gag cag gcg ccg ctc tct cta cca ttg acc ggg aat gga ctg      1152
291 Leu Gln Glu Gln Ala Pro Leu Ser Leu Pro Leu Thr Gly Asn Gly Leu
292      370      375      380
294 ttg tgt gta cgg gta gtg agc cgt gaa ggg gaa tac tat gag ttc aca      1200
295 Leu Cys Val Arg Val Val Ser Arg Glu Gly Glu Tyr Tyr Glu Phe Thr
296 385      390      395      400
298 ggt cga gcc gat agc cct cac ccc gta tcg ttt tga      1236
299 Gly Arg Ala Asp Ser Pro His Pro Val Ser Phe
300      405      410
303 <210> SEQ ID NO: 4
304 <211> LENGTH: 411
305 <212> TYPE: PRT
306 <213> ORGANISM: Unknown
308 <220> FEATURE:
309 <223> OTHER INFORMATION: Environmental DNA
311 <220> FEATURE:
312 <221> NAME/KEY: misc_feature
313 <222> LOCATION: (198)..(198)
314 <223> OTHER INFORMATION: Xaa = any amino acid
316 <220> FEATURE:
317 <221> NAME/KEY: misc_feature
318 <222> LOCATION: (594)..(594)
319 <223> OTHER INFORMATION: N= G, A, C or T
321 <400> SEQUENCE: 4
323 Met Pro Thr Leu Asp Trp Pro Gly Lys Gln Leu Ser Phe Pro Pro Ala
324 1      5      10      15
326 Thr Ser Leu His Leu Glu Ser Val Val Thr Glu Gly Ala Glu Ser Pro
327      20      25      30
329 Pro Asn Arg Leu Ile Trp Ala Asp Asn Leu Pro Leu Met Val Asp Leu
330      35      40      45
332 Leu Ala Glu Tyr Glu Gly Lys Ile Asp Leu Ile Tyr Ala Asp Pro Pro

```

← This is a protein sequence - amino acids only!
 R is not found at position 594

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/19/2002
PATENT APPLICATION: US/09/689,343C TIME: 14:59:22

Input Set : A:\NEB-181.ST25.txt
Output Set: N:\CRF3\07192002\I689343C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 594
Seq#:3; Xaa Pos. 198
Seq#:4; Xaa Pos. 198

VERIFICATION SUMMARY

DATE: 07/19/2002

PATENT APPLICATION: US/09/689,343C

TIME: 14:59:22

Input Set : A:\NEB-181.ST25.txt

Output Set: N:\CRF3\07192002\I689343C.raw

L:246 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:576
L:247 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:624
L:359 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:192